

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: HOSTETTER, MARGARET K.
GALE, CHERYL A.
BENDEL, CATHERINE M.
TAO, NIAN-JUN
KENDRICK, KATHLEEN

(ii) TITLE OF INVENTION: CANDIDA ALBICANS GENE, INTEGRIN-LIKE
PROTEIN, ANTIBODIES, AND METHODS OF USE

(iii) NUMBER OF SEQUENCES: 12

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: MUETING, RAASCH, GEBHARDT & SCHWAPPACH, P.A.
(B) STREET: 119 NORTH FOURTH STREET, SUITE 203
(C) CITY: MINNEAPOLIS
(D) STATE: MINNESOTA
(E) COUNTRY: USA
(F) ZIP: 55401

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/642,846
(B) FILING DATE: 03-MAY-1996
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: MUETING, ANN M.
(B) REGISTRATION NUMBER: 33,977
(C) REFERENCE/DOCKET NUMBER: 110.00280101

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 612-305-1217
(B) TELEFAX: 612-305-1228

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5194 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCCAAAAAAG ATAAATAAA AACAAAACAA AACAAAAGTA CTAACAAATT ATTGAACTT	60
TTAATTTTTA ATAAAGAATC AGTAGATCTA TTGTTAAAAG AAATGAACTC AACTCCAAGT	120

AAATTATTAC CGATAGATAA ACATTCTCAT TTACAATTAC AGCCTCAATC GTCCTCGGCA	180
TCAATATTTA ATTCCCCAAC AAAACCATTG AATTTCCCCA GAACAAATTC CAAGCCGAGT	240
TTAGATCCAA ATTCAAGCTC TGATACCTAC ACTAGCGAAC AAGATCAAGA GAAAGGGAAA	300
GAAGAGAAAA AGGACACAGC CTTTCAAACA TCTTTTGATA GAAATTTTGA TCTTGATAAT	360
TCAATCGATA TACAACAAAC AATTCAACAT CAGCAACAAC AGCCACAACA ACAACAACAA	420
CTCTCACAAA CCGACAATAA TTTAATTGAT GAATTTTCTT TTCAAACACC GATGACTTCG	480
ACTTTAGACC TAACCAAGCA AAATCCAAC TGGGACAAAG TGAATGAAAA TCATGCACCA	540
ACTTATATAA ATACCTCCCC CAACAAATCA ATAATGAAAA AGGCAACTCC TAAAGCGTCA	600
CCTAAAAAAG TTGCATTTAC TGTAACATAA CCCGAAATTC ATCATTATCC AGATAATAGA	660
GTCGAGGAAG AAGATCAAAG TCAACAAAAA GAAGATTGAG TTGAGCCACC CTTAATACAA	720
CATCAATGGA AAGATCCTTC TCAATTCAAT TATTCTGATG AAGATACAAA TGCTTCAGTT	780
CCACCAACAC CACCACTTCA TACGACGAAA CCTACTTTTG CGCAATTATT GAACAAAAAC	840
AACGAAGTCA ATCTGGAACC AGAGGCATTG ACAGATATGA AATTAAAGCG CGAAAATTTT	900
AGCAATTTAT CATTAGATGA AAAAGTCAAT TTATATCTTA GTCCCACTAA TAATAACAAT	960
AGTAAGAATG TGTCAGATAT GGATCTGCAT TTACAAAAC TGAAGACGC TTCGAAAAAC	1020
AAAACATAATG AAAATATTCA CAATTTGTCA TTTGCTTTAA AAGCACCAA GAATGATATT	1080
GAAAACCCAT TAAACTCATT GACTAACGCA GATATTCTGT TAAGATCATC TGGATCATCA	1140
CAATCGTCAT TACAATCTTT GAGGAATGAC AATCGTGTCT TGGAATCAGT GCCTGGGTCA	1200
CCTAAGAAGG TTAATCCTGG ATTGTCTTTG AATGACGGCA TAAAGGGGTT CTCTGATGAG	1260
GTTGTTGAAT CATTACTTCC TCGTGACTTA TCTCGAGACA AATTAGAGAC TACAAAAGAA	1320
CATGATGCAC CAGAACACAA CAATGAGAAT TTTATTGATG CTAAATCGAC TAATACCAAT	1380
AAGGGACAAC TCTTAGTATC ATCTGATGAT CATTTGGACT CTTTGTATAG ATCCTATAAC	1440
CACACTGAAC AATCAATTTT GAATCTTTTG AATAGTGCAT CACAATCTCA AATTTGTTA	1500
AATGCATTGG AAAAACAAG GCAACACAG GAACAAGAAC AAACACAAGC GGCAGAGCCT	1560
GAAGAAGAAA CTTGTTTTAG TGATAATATC AAAGTTAAAC AAGAGCCAAA GAGCAATTTG	1620
GAGTTTGTCA AGGTTACCAT CAAGAAAGAA CCAGTTCTGG CCACGGAAAT AAAAGCTCCA	1680
AAAAGAGAAT TTTCAAGTCG AATATTAAGA ATAAAAATG AAGATGAAAT TGCCGAACCA	1740
GCTGATATTC ATCCTAAAAA AGAAAATGAA GCAAACAGTC ATGTCGAAGA TACTGATGCA	1800
TTGTTGAAGA AAGCACTTAA TGATGATGAG GAATCTGACA CGACCCAAAA CTCAACGAAA	1860
ATGTCAATTC GTTTTCATAT TGATAGTGAT TGGAAATTGG AAGACAGTAA TGATGGCGAT	1920
AGAGAAGATA ATGATGATAT TTCTCGTTTT GAGAAATCAG ATATTTTGAA CGACGTATCA	1980
CAGACTTCTG ATATTATTGG TGACAAATAT GGAACTCAT CAAGTGAAAT AACCACCAAA	2040

ACATTAGCAC	CCCCAAGATC	GGACAACAAT	GACAAGGAGA	ATTCTAAATC	TTTGGAAGAT	2100
CCAGCTAATA	ATGAATCATT	GCAACAACAA	TTGGAGGTAC	CGCATACAAA	AGAAGATGAT	2160
AGCATTTTAG	CCAACTCGTC	CAATATTGCT	CCACCTGAAG	AATTGACTTT	GCCCGTAGTG	2220
GAAGCAAATG	ATTATTCATC	TTTTAATGAC	GTGACCAAAA	CTTTTGATGC	ATACTCAAGC	2280
TTTGAAGAGT	CATTATCTAG	AGAGCACGAA	ACTGATTCAA	AACCAATTAA	TTTCATATCA	2340
ATTTGGCATA	AACAAGAAAA	GCAGAAGAAA	CATCAAATTC	ATAAAGTTCC	AACTAAACAG	2400
ATCATTGCTA	GTTATCAACA	ATACAAAAAC	GAACAAGAAT	CTCGTGTTAC	TAGTGATAAA	2460
GTGAAAATCC	CAAATGCCAT	ACAATTCAAG	AAATTCAAAG	AGGTAAATGT	CATGTCAAGA	2520
AGAGTTGTTA	GTCCAGACAT	GGATGATTTG	AATGTATCTC	AATTTTTACC	AGAATTATCT	2580
GAAGACTCTG	GATTTAAAGA	TTTGAATTTT	GCCAACTACT	CCAATAACAC	CAACAGACCA	2640
AGAAGTTTTA	CTCCATTGAG	CACTAAAAAT	GTCTTGTCGA	ATATTGATAA	CGATCCTAAT	2700
GTTGTTGAAC	CTCCTGAACC	GAAATCATAT	GCTGAAATTA	GAAATGCTAG	ACGGTTATCA	2760
GCTAATAAGG	CAGCGCCAAA	TCAGGCACCA	CCATTGCCAC	CACAACGACA	ACCATCTTCA	2820
ACTCGTTCCA	ATTCAAATAA	ACGAGTGTCC	AGATTTAGAG	TGCCCACATT	TGAAATTAGA	2880
AGAACTTCTT	CAGCATTAGC	ACCTTGTGAC	ATGTATAATG	ATATTTTTGA	TGATTTCTGGT	2940
GCGGGTTCTA	AACCAACTAT	AAAGGCAGAA	GGAATGAAAA	CATTGCCAAG	TATGGATAAA	3000
GATGATGTCA	AGAGGATTTT	GAATGCAAAG	AAAGGTGTGA	CTCAAGATGA	ATATATAAAT	3060
GCCAAACTTG	TTGATCAAAA	ACCTAAAAAG	AATTCAATTG	TCACCGATCC	CGAAGACCGA	3120
TATGAAGAAT	TACAACAAAC	TGCCTCTATA	CACAATGCCA	CCATTGATTC	AAGTATTTAT	3180
GGCCGACCAG	ACTCCATTTT	TACCGACATG	TTGCCTTATC	TTAGTGATGA	ATTGAAAAAA	3240
CCACCTACGG	CTTTATTATC	TGCTGATCGT	TTGTTTATGG	AACAAGAAGT	ACATCCGTTA	3300
AGATCAAAC	CTGTTTTGGT	TCACCCAGGG	GCAGGAGCAG	CAACTAATTC	TTCAATGTTA	3360
CCAGAGCCAG	ATTTTGAATT	AATCAATTCA	CCTGCTAGAA	ATGTGCTGAA	CAACAGTGAT	3420
AATGTCGCCA	TCAGTGGTAA	TGCTAGTACT	ATTAGTTTTA	ACCAATTGGA	TATGAATTTT	3480
GATGACCAAG	CTACAATTGG	TCAAAAAATC	CAAGAGCAAC	CTGCTTCAA	ATCCGCCAAT	3540
ACTGTTCTGT	GTGATGATGA	TGGATTGGCC	AGTGCACCTG	AAACACCAAG	AACTCCTACC	3600
AAAAAGGAGT	CCATATCAAG	CAAGCCTGCC	AAGCTTTCTT	CTGCCTCCCC	TAGAAAAATCA	3660
CCAATTAAGA	TTGGTTCACC	AGTTCGAGTT	ATTAAGAAAA	ATGGATCAAT	TGCTGGCATT	3720
GAACCAATCC	CAAAAGCCAC	TCACAAACCG	AAGAAATCAT	TCCAAGGAAA	CGAGATTTCA	3780
AACCATAAAG	TACGAGATGG	TGGAATTTCA	CCAAGCTCCG	GATCAGAGCA	TCAACAGCAT	3840
AATCCTAGTA	TGGTTTCTGT	TCCTTCACAG	TATACTGATG	CTACTTCAAC	GGTTCCAGAT	3900
GAAAACAAAG	ATGTTCAACA	CAAGCCTCGT	GAAAAGCAAA	AGCAAAAGCA	TCACCATCGC	3960

CATCATCATC ATCATCATAA ACAAAAAA CT GATATTCCGG GTGTTGTTGA TGATGAAATT 4020
 CCTGATGTAG GATTACAAGA ACGAGGCAAA TTATTCTTTA GAGTTTTAGG AATTAAGAAT 4080
 ATCAATTTAC CCGATATTAA TACTCACAAA GGAAGATTCA CTTTAACGTT GGATAATGGA 4140
 GTGCATTGTG TTA CTACACC AGAATACAAC ATGGACGACC ATAATGTTGC CATAGGTAAA 4200
 GAATTTGAGT TGACAGTTGC TGATTCATTA GAGTTTATTT TAACTTTGAA GGCATCATAT 4260
 GAAAAACCTC GTGGTACATT AGTAGAAGTG ACTGAAAAGA AAGTTGTCAA ATCAAGAAAT 4320
 AGATTGAGTC GATTATTTGG ATCGAAAGAT ATTATCACCA CGACAAAGTT TGTGCCCACT 4380
 GAAGTCAAAG ATACCTGGGC TAATAAGTTT GCTCCTGATG GTTCATTTGC TAGATGTTAC 4440
 ATTGATTTAC AACAATTTGA AGACCAAATC ACCGGTAAAG CATCACAGTT TGATCTCAAT 4500
 TGTTTTAATG AATGGGAAAC TATGAGTAAT GGCAATCAAC CAATGAAAAG AGGCAAACCT 4560
 TATAAGATTG CTCAATTGGA AGTTAAAATG TTGTATGTTT CACGATCAGA TCCAAGAGAA 4620
 ATATTACCAA CCAGCATTAG ATCCGCATAT GAAAGCATCA ATGAATTTAA CAATGAACAG 4680
 AATAATTACT TTGAAGGTTA TTTACATCAA GAAGGAGGTG ATTGTCCAAT TTTTAAGAAA 4740
 CGTTTTTTCA AATTAATGGG CACTTCTTTA TTGGCTCATA GTGAAATATC TCATAAAACT 4800
 AGAGCCAAAA TTAATTTATC AAAAGTTGTT GATTTGATTT ATGTTGATAA AGAAAACATT 4860
 GATCGTTCCA ATCATCGAAA TTTCAGTGAT GTGTTATTGT TGGATCATGC ATTCAAATC 4920
 AAATTTGCTA ATGGTGAGTT GATTGATTTT TGTGCTCCTA ATAAACATGA AATGAAAATA 4980
 TGGATTCAAA ATTTACAAGA AATTATCTAT AGAAATCGGT TCAGACGTCA ACCATGGGTA 5040
 AATTTGATGC TTCAACAACA ACAACAACA CAACAACAAC AAAGCTCCCA ACAGTAATTG 5100
 AAAGGTCTAC TTTTGATTTT TTTAATTTTA ATTGGCAAAT ATATGCCCAT TTTGTATTAT 5160
 CTTTGTAGTCT AATAGCGTTT TCTTTTTTTC CAGT 5194

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1664 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Asn Ser Thr Pro Ser Lys Leu Leu Pro Ile Asp Lys His Ser His
 1 5 10 15
 Leu Gln Leu Gln Pro Gln Ser Ser Ser Ala Ser Ile Phe Asn Ser Pro
 20 25 30
 Thr Lys Pro Leu Asn Phe Pro Arg Thr Asn Ser Lys Pro Ser Leu Asp
 4

	35					40					45				
Pro	Asn 50	Ser	Ser	Ser	Asp	Thr 55	Tyr	Thr	Ser	Glu	Gln 60	Asp	Gln	Glu	Lys
Gly 65	Lys	Glu	Glu	Lys	Lys 70	Asp	Thr	Ala	Phe	Gln 75	Thr	Ser	Phe	Asp	Arg 80
Asn	Phe	Asp	Leu	Asp 85	Asn	Ser	Ile	Asp	Ile 90	Gln	Gln	Thr	Ile	Gln 95	His
Gln	Gln	Gln	Gln 100	Pro	Gln	Gln	Gln	Gln 105	Gln	Leu	Ser	Gln	Thr 110	Asp	Asn
Asn	Leu	Ile 115	Asp	Glu	Phe	Ser	Phe 120	Gln	Thr	Pro	Met	Thr 125	Ser	Thr	Leu
Asp	Leu 130	Thr	Lys	Gln	Asn	Pro 135	Thr	Val	Asp	Lys	Val 140	Asn	Glu	Asn	His
Ala 145	Pro	Thr	Tyr	Ile	Asn 150	Thr	Ser	Pro	Asn	Lys 155	Ser	Ile	Met	Lys	Lys 160
Ala	Thr	Pro	Lys	Ala 165	Ser	Pro	Lys	Lys	Val 170	Ala	Phe	Thr	Val	Thr 175	Asn
Pro	Glu	Ile	His 180	His	Tyr	Pro	Asp	Asn 185	Arg	Val	Glu	Glu	Glu 190	Asp	Gln
Ser	Gln	Gln 195	Lys	Glu	Asp	Ser	Val 200	Glu	Pro	Pro	Leu	Ile 205	Gln	His	Gln
Trp	Lys 210	Asp	Pro	Ser	Gln	Phe 215	Asn	Tyr	Ser	Asp	Glu 220	Asp	Thr	Asn	Ala
Ser 225	Val	Pro	Pro	Thr	Pro 230	Pro	Leu	His	Thr	Thr 235	Lys	Pro	Thr	Phe	Ala 240
Gln	Leu	Leu	Asn	Lys 245	Asn	Asn	Glu	Val	Asn 250	Ser	Glu	Pro	Glu	Ala 255	Leu
Thr	Asp	Met	Lys 260	Leu	Lys	Arg	Glu	Asn 265	Phe	Ser	Asn	Leu	Ser 270	Leu	Asp
Glu	Lys	Val 275	Asn	Leu	Tyr	Leu	Ser 280	Pro	Thr	Asn	Asn	Asn 285	Asn	Ser	Lys
Asn	Val 290	Ser	Asp	Met	Asp	Ser 295	His	Leu	Gln	Asn	Leu 300	Gln	Asp	Ala	Ser
Lys 305	Asn	Lys	Thr	Asn	Glu 310	Asn	Ile	His	Asn	Leu 315	Ser	Phe	Ala	Leu	Lys 320
Ala	Pro	Lys	Asn	Asp 325	Ile	Glu	Asn	Pro	Leu 330	Asn	Ser	Leu	Thr	Asn 335	Ala
Asp	Ile	Ser	Leu 340	Arg	Ser	Ser	Gly	Ser 345	Ser	Gln	Ser	Ser	Leu 350	Gln	Ser
Leu	Arg	Asn 355	Asp	Asn	Arg	Val	Leu 360	Glu	Ser	Val	Pro	Gly 365	Ser	Pro	Lys
Lys	Val 370	Asn	Pro	Gly	Leu	Ser 375	Leu	Asn	Asp	Gly	Ile 380	Lys	Gly	Phe	Ser

Phe Asp Ala Tyr Ser Ser Phe Glu Glu Ser Leu Ser Arg Glu His Glu
725 730 735
Thr Asp Ser Lys Pro Ile Asn Phe Ile Ser Ile Trp His Lys Gln Glu
740 745 750
Lys Gln Lys Lys His Gln Ile His Lys Val Pro Thr Lys Gln Ile Ile
755 760 765
Ala Ser Tyr Gln Gln Tyr Lys Asn Glu Gln Glu Ser Arg Val Thr Ser
770 775 780
Asp Lys Val Lys Ile Pro Asn Ala Ile Gln Phe Lys Lys Phe Lys Glu
785 790 795 800
Val Asn Val Met Ser Arg Arg Val Val Ser Pro Asp Met Asp Asp Leu
805 810 815
Asn Val Ser Gln Phe Leu Pro Glu Leu Ser Glu Asp Ser Gly Phe Lys
820 825 830
Asp Leu Asn Phe Ala Asn Tyr Ser Asn Asn Thr Asn Arg Pro Arg Ser
835 840 845
Phe Thr Pro Leu Ser Thr Lys Asn Val Leu Ser Asn Ile Asp Asn Asp
850 855 860
Pro Asn Val Val Glu Pro Pro Glu Pro Lys Ser Tyr Ala Glu Ile Arg
865 870 875 880
Asn Ala Arg Arg Leu Ser Ala Asn Lys Ala Ala Pro Asn Gln Ala Pro
885 890 895
Pro Leu Pro Pro Gln Arg Gln Pro Ser Ser Thr Arg Ser Asn Ser Asn
900 905 910
Lys Arg Val Ser Arg Phe Arg Val Pro Thr Phe Glu Ile Arg Arg Thr
915 920 925
Ser Ser Ala Leu Ala Pro Cys Asp Met Tyr Asn Asp Ile Phe Asp Asp
930 935 940
Phe Gly Ala Gly Ser Lys Pro Thr Ile Lys Ala Glu Gly Met Lys Thr
945 950 955 960
Leu Pro Ser Met Asp Lys Asp Asp Val Lys Arg Ile Leu Asn Ala Lys
965 970 975
Lys Gly Val Thr Gln Asp Glu Tyr Ile Asn Ala Lys Leu Val Asp Gln
980 985 990
Lys Pro Lys Lys Asn Ser Ile Val Thr Asp Pro Glu Asp Arg Tyr Glu
995 1000 1005
Glu Leu Gln Gln Thr Ala Ser Ile His Asn Ala Thr Ile Asp Ser Ser
1010 1015 1020
Ile Tyr Gly Arg Pro Asp Ser Ile Ser Thr Asp Met Leu Pro Tyr Leu
1025 1030 1035 1040
Ser Asp Glu Leu Lys Lys Pro Pro Thr Ala Leu Leu Ser Ala Asp Arg
1045 1050 1055
Leu Phe Met Glu Gln Glu Val His Pro Leu Arg Ser Asn Ser Val Leu

1060	1065	1070
Val His Pro Gly Ala Gly Ala	Ala Thr Asn Ser Ser	Met Leu Pro Glu
1075	1080	1085
Pro Asp Phe Glu Leu Ile	Asn Ser Pro Ala Arg	Asn Val Ser Asn Asn
1090	1095	1100
Ser Asp Asn Val Ala Ile	Ser Gly Asn Ala Ser Thr Ile	Ser Phe Asn
1105	1110	1115
Gln Leu Asp Met Asn Phe Asp Asp	Gln Ala Thr Ile Gly Gln Lys Ile	
1125	1130	1135
Gln Glu Gln Pro Ala Ser Lys Ser	Ala Asn Thr Val Arg Gly Asp Asp	
1140	1145	1150
Asp Gly Leu Ala Ser Ala Pro	Glu Thr Pro Arg Thr Pro Thr Lys Lys	
1155	1160	1165
Glu Ser Ile Ser Ser Lys Pro Ala Lys Leu	Ser Ser Ala Ser Pro Arg	
1170	1175	1180
Lys Ser Pro Ile Lys Ile Gly Ser Pro Val	Arg Val Ile Lys Lys Asn	
1185	1190	1195
Gly Ser Ile Ala Gly Ile Glu Pro Ile	Pro Lys Ala Thr His Lys Pro	
1205	1210	1215
Lys Lys Ser Phe Gln Gly Asn Glu Ile	Ser Asn His Lys Val Arg Asp	
1220	1225	1230
Gly Gly Ile Ser Pro Ser Ser Gly Ser Glu His	Gln Gln His Asn Pro	
1235	1240	1245
Ser Met Val Ser Val Pro Ser Gln Tyr Thr	Asp Ala Thr Ser Thr Val	
1250	1255	1260
Pro Asp Glu Asn Lys Asp Val Gln His Lys	Pro Arg Glu Lys Gln Lys	
1265	1270	1275
Gln Lys His His His Arg His His His	His His His Lys Gln Lys Thr	
1285	1290	1295
Asp Ile Pro Gly Val Val Asp Asp Glu Ile	Pro Asp Val Gly Leu Gln	
1300	1305	1310
Glu Arg Gly Lys Leu Phe Phe Arg Val Leu	Gly Ile Lys Asn Ile Asn	
1315	1320	1325
Leu Pro Asp Ile Asn Thr His Lys Gly Arg	Phe Thr Leu Thr Leu Asp	
1330	1335	1340
Asn Gly Val His Cys Val Thr Thr Pro Glu	Tyr Asn Met Asp Asp His	
1345	1350	1355
Asn Val Ala Ile Gly Lys Glu Phe Glu Leu	Thr Val Ala Asp Ser Leu	
1365	1370	1375
Glu Phe Ile Leu Thr Leu Lys Ala Ser Tyr	Glu Lys Pro Arg Gly Thr	
1380	1385	1390
Leu Val Glu Val Thr Glu Lys Lys Val Val	Lys Ser Arg Asn Arg Leu	
1395	1400	1405

Ser Arg Leu Phe Gly Ser Lys Asp Ile Ile Thr Thr Thr Lys Phe Val
 1410 1415 1420
 Pro Thr Glu Val Lys Asp Thr Trp Ala Asn Lys Phe Ala Pro Asp Gly
 1425 1430 1435 1440
 Ser Phe Ala Arg Cys Tyr Ile Asp Leu Gln Gln Phe Glu Asp Gln Ile
 1445 1450 1455
 Thr Gly Lys Ala Ser Gln Phe Asp Leu Asn Cys Phe Asn Glu Trp Glu
 1460 1465 1470
 Thr Met Ser Asn Gly Asn Gln Pro Met Lys Arg Gly Lys Pro Tyr Lys
 1475 1480 1485
 Ile Ala Gln Leu Glu Val Lys Met Leu Tyr Val Pro Arg Ser Asp Pro
 1490 1495 1500
 Arg Glu Ile Leu Pro Thr Ser Ile Arg Ser Ala Tyr Glu Ser Ile Asn
 1505 1510 1515 1520
 Glu Leu Asn Asn Glu Gln Asn Asn Tyr Phe Glu Gly Tyr Leu His Gln
 1525 1530 1535
 Glu Gly Gly Asp Cys Pro Ile Phe Lys Lys Arg Phe Phe Lys Leu Met
 1540 1545 1550
 Gly Thr Ser Leu Leu Ala His Ser Glu Ile Ser His Lys Thr Arg Ala
 1555 1560 1565
 Lys Ile Asn Leu Ser Lys Val Val Asp Leu Ile Tyr Val Asp Lys Glu
 1570 1575 1580
 Asn Ile Asp Arg Ser Asn His Arg Asn Phe Ser Asp Val Leu Leu Leu
 1585 1590 1595 1600
 Asp His Ala Phe Lys Ile Lys Phe Ala Asn Gly Glu Leu Ile Asp Phe
 1605 1610 1615
 Cys Ala Pro Asn Lys His Glu Met Lys Ile Trp Ile Gln Asn Leu Gln
 1620 1625 1630
 Glu Ile Ile Tyr Arg Asn Arg Phe Arg Arg Gln Pro Trp Val Asn Leu
 1635 1640 1645
 Met Leu Gln Gln Gln Gln Gln Gln Gln Gln Gln Ser Ser Gln Gln
 1650 1655 1660

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 236 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (D) OTHER INFORMATION: amino acid positions 218-453 from SEQ ID NO:2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ser Asp Glu Asp Thr Asn Ala Ser Val Pro Pro Thr Pro Pro Leu His
 1 5 10 15
 Thr Thr Lys Pro Thr Phe Ala Gln Leu Leu Asn Lys Asn Asn Glu Val
 20 25 30
 Asn Ser Glu Pro Glu Ala Leu Thr Asp Met Lys Leu Lys Arg Glu Asn
 35 40 45
 Phe Ser Asn Leu Ser Leu Asp Glu Lys Val Asn Leu Tyr Leu Ser Pro
 50 55 60
 Thr Asn Asn Asn Asn Ser Lys Asn Val Ser Asp Met Asp Ser His Leu
 65 70 75 80
 Gln Asn Leu Gln Asp Ala Ser Lys Asn Lys Thr Asn Glu Asn Ile His
 85 90 95
 Asn Leu Ser Phe Ala Leu Lys Ala Pro Lys Asn Asp Ile Glu Asn Pro
 100 105 110
 Leu Asn Ser Leu Thr Asn Ala Asp Ile Ser Leu Arg Ser Ser Gly Ser
 115 120 125
 Ser Gln Ser Ser Leu Gln Ser Leu Arg Asn Asp Asn Arg Val Leu Glu
 130 135 140
 Ser Val Pro Gly Ser Pro Lys Lys Val Asn Pro Gly Leu Ser Leu Asn
 145 150 155 160
 Asp Gly Ile Lys Gly Phe Ser Asp Glu Val Val Glu Ser Leu Leu Pro
 165 170 175
 Arg Asp Leu Ser Arg Asp Lys Leu Glu Thr Thr Lys Glu His Asp Ala
 180 185 190
 Pro Glu His Asn Asn Glu Asn Phe Ile Asp Ala Lys Ser Thr Asn Thr
 195 200 205
 Asn Lys Gly Gln Leu Leu Val Ser Ser Asp Asp His Leu Asp Ser Phe
 210 215 220
 Asp Arg Ser Tyr Asn His Thr Glu Gln Ser Ile Leu
 225 230 235

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 23 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Tyr Leu Ser Pro Thr Asn Asn Asn Asn Ser Lys Asn Val Ser Asp Met
 1 5 10 15
 10

Asp Leu His Leu Gln Asn Leu
20

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Asp Trp Lys Leu Glu Asp Ser Asn Asp Gly Asp Arg Glu Asp Asn Asp
1 5 10 15

Asp Ile Ser Arg Phe Glu Lys
20

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Ser Lys Ser Ala Asn Thr Val Arg Gly Asp Asp Asp Gly Leu Ala Ser
1 5 10 15

Ala

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Asp His Leu Asp Ser Phe Asp Arg Ser Tyr Asn His Thr Glu Gln Ser
1 5 10 15

Ile

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Trp Ile Gln Asn Leu Gln Glu Ile Ile Tyr Arg Asn Arg Phe Arg Arg
1 5 10 15
Gln

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GAATTCAATG CTACCCTCAA

20

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CCCGGGGGAC CCCCTTCACT

20

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AARGTYGGWT TYTTYAAR

18

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GAAATHGAYG AYTTRATG

18